视觉统计概要表征的机制

张秀玲[1**]　张帆[1**]　葛明晓[1]　李思嘉[1]　蒋毅[2*]
[1]（东北师范大学心理学院　长春 130024）
[2]（中国科学院心理研究所，脑与认知科学国家重点实验室　北京 100101）

摘要：尽管感觉登记是容量无限的，但不意味着视觉系统无需高效地处理我们日常面对的海量的视觉信息，统计概要表征便是一种高效的信息加工过程，当呈现由多个视觉刺激构成的集合，视觉系统可以快速提取集合的平均属性等统计概要信息。统计概要表征是理解神经计算过程的重要心理过程。目前有很多研究揭示了这一过程的某些认知机制，如自动性加工机制，不同水平或特征间的特异性或通用性的机制。然而仅有少量研究关注其神经机制，未来需要更多的研究采用神经科学技术手段直接考察统计概要表征的神经机制。
关键词：统计概要表征；集合编码；整体加工；神经机制
分类号：B842

Mechanism of Visual Statistical Summary Representations

Zhang Xiuling[1**]　Zhang Fan[1**]　Ge Mingxiao[1]　Li Sijia[1]　Jiang Yi[2*]
[1](School of Psychology, Northeast Normal University, Changchun 130024, China)
[2](State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China)

Abstract: Despite the unlimited capacity of sensory registration, visual system can still provide an efficient summary of a cluttered scene, representing the statistical properties of multiple objects rather than forming detailed representations of individual objects. There is growing interest in the behavioral study of statistical summary representations (SSRs), especially in the exploring of their automatic mechanism as well as the domain-general or domain specific ensemble mechanism. However, the neural underpinnings of SSRs have received far less attention. Future work on SSRs may use neuroimaging methods to investigate their neural substrates directly, which is also important for understanding neural computation.
Keywords: statistical summary representation; ensemble coding; ensemble processing; neural mechanism

1 引言

感觉登记是容量无限的，这似乎意味着我们的视觉系统无需考虑加工效率的问题，可交由注意和工作记忆去处理。事实上，我们的视觉系统不但能够进行知觉组织，还能够从集群中提取出核心关键信息。当呈现包含一组相似刺激的集合时，视觉系统能够快速准确地提取集合的整体统计属性(而不是单个刺激属性)，形成统计概要表征(Ariely, 2001; Haberman & Whitney, 2007)。统计概要表征和知觉组织都反映了视觉系统加工的效率性。

统计概要表征是研究神经计算的重要心理过程。在对单个刺激的编码基础上，神经元形成了对多个刺激构成的集合的统计概要表征，这背后的神经计算

** 为共同第一作者及共同通讯作者
* 为共同通讯作者

过程是怎样的，其特点是什么，这些问题吸引了大量研究者的关注。本综述涵盖了关于统计概要表征自动性加工机制的内容，以及关于其特异性和通用性机制的争论，重点介绍了近年来采用认知神经科学手段研究其神经机制的研究。并在此基础上进行总结与展望，为之后开展统计概要表征神经机制的研究提供思路与方向。

集中趋势(如平均值)和离散趋势(如方差)是视觉统计概要表征的两个重要的指标(Haberman, lee, et al., 2015; Norman et al., 2015)。对于一个刺激集合，想要对其进行较完整的表征，集中趋势表征和离散趋势表征似乎是缺一不可(全可 等, 2015)。平均表征是描述刺激集合集中趋势的一个重要代表。Ariely (2001)最初对圆的大小统计概要表征的研究中发现平均表征快速、准确，且不受集合大小即成员数目的影响，但不能够对集合内的个体进行准确表征，说明集合平均表征比成员表征更占优势。除了刺激的大小，研究者在从低水平到高水平的很多种刺激中发现了平均表征，如刺激大小(Haberman & Suresh, 2021)、空间位置(Sun et al., 2021)、深度(Wardle et al., 2012)、朝向(Parkes et al., 2001)、运动方向(Watamaniuk & Duchon, 1992)、颜色和对比度(Rajendran et al., 2021)、亮度(Bauer, 2009)以及更高水平的面孔情绪(Haberman, Harp, et al., 2009)、面孔性别(Haberman & Whitney, 2007)、面孔身份(de Fockert & Wolfenstein, 2009; Davis et al., 2021)、目光朝向(Florey et al., 2016)和生物运动(Sweeny et al., 2013; Nguyen et al., 2021) 等。

方差表征是描述刺激集合离散趋势的一个重要代表。Haberman, Lee 和 Whitney (2015)首次从方差角度探讨了面孔情绪的统计概要表征。实验采用调整匹配法，要求被试根据一组情绪不同的面孔集合的方差，调整另一组情绪面孔集合的方差与之匹配，并以被试匹配的平均误差为指标，结果发现被试可以很好地表征情绪面孔集合的方差，且不受集合大小的影响。之后很多研究均发现我们的视觉系统可以计算各种视觉特征的方差，例如颜色(Maule & Franklin, 2020; Ward et al., 2016)、亮度(Tong et al., 2015)、大小(Tokita et al., 2016)、面孔性别和种族(Phillips et al., 2018)，这说明视觉系统对集合的离散趋势可以进行很好的表征。

2 视觉统计概要表征的整体自动性加工

全可等人(2015)的综述介绍了统计概要表征的两种加工方式，基于整体的平行分布式加工(不受集合大小即成员数量的影响)，以及基于个体的抽样加工。整体的加工不同于个体加工的机制，被试能够准确表征集合平均值，却不能够对集合内的个体进行准确表征(Ariely, 2001; Whitney & Yamanashi Leib, 2018)。整体的加工是自动的(“快速，任务无关刺激也可以形成统计表征，不受注意分配的影响，即使不能检测局部变化也能检测整体变化”)(全可等, 2015)。 因此在本综述中很多关于加工时间和注意资源的研究不再赘述，仅介绍新的神经科学证据及无意识研究证据。

自动的加工有以下特征：快速、无意识的、不受注意资源的影响(分布式注意或平行加工，在注意资源较少时依然可以得到加工)、非随意的(unintentional)、不可控的。 需要注意的是，满足所有自动特征的认知过程和行为非常少，甚至在驾驶活动中也不能满足所有以上特征，因此在理解和区分自动-非自动过程时，并不需要教条地采用全或无的标准。

值得高兴的是统计概要表征的自动性加工已有少量的神经科学研究提供了支持性证据。首先在注意的影响上，一项 ERP 研究探讨了多个面孔表情的平均表征是否需要注意，使用线索提示范式，要求被试在有效线索以及无效线索条件下判断面孔集合的平均表情或者单个面孔的表情是积极还是消极的，结果发现在无效线索条件下，被试依然可以提取平均表情，同时并没有观察到显著的 N2pc 成分(该成分反映了对目标刺激的空间选择性注意)，这表明被试可以在有限注意力条件下提取平均表情。有趣的是，有效线索条件下，集合任务和个体任务的 SPCN 成分(通常随着视觉短时记忆中项目数量的增加而增加，直至被试的记忆容量)没有显著差异，意味着集合中的多个面孔可能被压缩成一个“单一”对象存储到视觉短期记忆中(Ji et al., 2018)。

其次，在加工速度或加工时间上，一项 ERP 的研究采用偏差刺激(oddball)范式探究了由多个线条构成的集合表征的时间进程。研究者区分了基于集合中某个个体的 oddball 和基于整个集合的 oddball，发现与个体相比，基于整体的 oddball 诱发的 P3b 潜伏期明显更早。多变量模式分析(MVPA)也发现，和个体条件相比，神经信号能够更早地区分整体条件下的标准刺激和偏差刺激(分类开始于 102ms)。这表明整体感知可以快速发生，且对整体的表征早于对成员单个对象属性的表征(Epstein & Emmanouil, 2021)，这与 Li 等人(2016)行为研究结果一致，发现在 50ms 的时候，平均表征就已经开始且好于成员表征。需要指出的是虽然整体加工比成员加工更早，一些行为研究揭示出统计概要表征发生于相对较晚的阶段，至少比大小错觉、利用深度线索的知觉恒常性以及双眼融合更晚。具体如下，平均大小判断受艾宾浩斯错觉的影响，表明平均大小可能是在知觉大小之后计算的(Im & Chong, 2009)；在双眼视差的深度线索下，统计概要表征会依据物体的真实大小进行表征，而非视网膜大小，即在知觉恒常性的基础上进行平均表征(Tiurina & Utochkin, 2019)；双眼竞争的研究也发现统计概要表征的形成发生于双眼融合之后(Joo et al., 2009)。

无意识对统计概要表征的影响目前存在争议。采用无意识研究中的拥挤范式(Crowding，即当一个位于外周视野的目标物体周围有其它物体呈现时，对这个目标物体的辨别会变得困难)， 发现尽管被试不能报告单个光栅朝向，但却可以可靠地估计光栅集合的平均朝向，表明在“Crowding”下，初级视觉皮层的朝向信息虽还未到达意识但已经得到平均表征(Parkes et al., 2001)。类似地，对位置以及面孔情绪的平均表征可以突破“Crowding”瓶颈(Fischer & Whitney, 2011; Greenwood et al., 2009)，但是在圆的大小平均上，无论是拥挤范式还是双眼竞争，无意识的成员个体都不能参与平均表征(Banno & Saiki, 2012; Joo et al., 2009)。

3 视觉统计概要表征的编码机制——特异性还是通用性？

3.1 不同水平及不同特征间统计概要表征的编码机制

不同视觉水平或不同特征之间到底存在各自分离的神经编码还是通用的神经编码，即视觉统计概要表征的神经机制是领域特异性还是通用性（或普遍性、一般性）的问题。Whitney 等人(2014)的观点认为视觉平均表征可能在视觉系统中的多个水平上进行表征。一些集合，如平均亮度、颜色和朝向，可能在最早的皮层(甚至可能是皮层下)产生，而面孔平均身份等可能在腹侧通路的后期阶段进行表征。

Haberman 等人(2015)首先通过行为实验对不同水平的视觉刺激具有编码特

异性的观点进行了考察。他们采用个体差异法，让观察者提取了不同刺激集合的平均特征，在一系列实验中(面孔身份和光栅朝向、面孔情绪和圆点颜色、三角形朝向和三角形颜色、光栅朝向和圆点颜色、光栅朝向和三角形朝向、面孔身份和面孔情绪)，研究者计算被试对两种刺激平均表征的误差相关性，以考察不同领域的统计概要表征是否存在一个共同机制或多个机制，结果发现个体在高-低水平领域间平均表征成绩不相关，认为存在特异于高-低不同水平的多重表征机制，即高-低水平领域间的统计概要表征具有特异性，而非跨越不同水平的单一的、通用的表征机制。这得到了后续一些研究的支持。Peng 等人(2019)发现互依型自我构念的启动增强了面孔身份的平均表征，但对圆点大小的平均表征没有影响。Sama 等人(2019)使用面孔集合刺激，通过一系列的行为实验，揭示了高水平特征(面孔身份)和低水平特征(面孔视角)的平均表征的独立性。

值得注意的是，虽然 Haberman 等人(2015)发现高和低水平各自领域内不同特征间平均表征成绩相关，但是并没有下结论同一水平不同特征间的平均表征是通用机制还是不同机制，认为存在两种可能性，一是同一水平领域内的统计概要表征确实共有一个通用性机制，二是可能存在误差导致不同特征间平均表征结果相关。然而另有些研究根据相关的结果得出了领域内涉及通用性机制的结论。事实上，关于高-低各自水平内不同特征间是否存在通用性机制，不同实证研究间也确实存在一些争议。

关于低水平领域内的研究，多数研究认为不同特征的统计概要表征是特异性的。Rajendran 等人(2021)对颜色和亮度平均表征的研究中，发现颜色和亮度平均表征存在差异，颜色平均表征需要间接推断，而亮度可以直接进行统计估计，可能反映了两种概要表征机制的特异性。Attarha 和 Moore (2015)以光栅为刺激材料，发现被试可以同时表征光栅的大小和朝向，说明大小和朝向的平均表征不受加工容量的限制，也暗示大小和朝向的统计概要表征的机制是独立的。Yüorük 和 Boduroglu (2020)的研究表明线段长度以及朝向间平均表征不相关，具有各自独立的机制。而 Kacin 等人(2021)在尝试验证 Yüorük 和 Boduroglu 的研究却得出了相反结论，认为不同特征间存在共同的即通用性机制，他们改进了刺激材料参数的设置以及任务的设置，发现线段长度和朝向的平均表征存在相关。类似地，Yang 等人(2018)使用个体差异法发现草莓大小的平均表征与棒棒糖朝向的平均表征的正确率相关显著，研究者认为大小和朝向的概要表征共享一个机制，他们的观点是，如果某两个特征的平均表征的成绩呈强相关，则说明这两个特征的平均表征存在一个共同的机制。

高水平领域内不同特征间也存在统计概要表征特异性。 Haberman 和 Ulrich (2019)发现面孔身份和面孔表情的平均表征精度存在一定差异。 在面孔倒置的研究中也发现了平均身份和平均表情的分离，Haberman 和 Whitney (2009)在面孔情绪平均表征研究中发现倒置降低了情绪平均表征的精度，而另一些研究发现倒置不影响面孔身份集合的平均表征(Sun & Choo, 2020; Davis et al., 2021)。平均身份和平均情绪的差异也许基于身份和情绪本身的差异，Haxby 等人(2000)在关于“面孔知觉的分布式人类神经系统”模型中提到了面孔身份和面孔表情加工编码的差异，其中梭状回主要分析面孔不变部分的特征，在身份识别中起重要作用，而颞上沟主要分析面孔可变部分的特征，在感知面孔表情、眼睛注视方向和嘴唇运动中起重要作用。

目前对不同领域的方差表征是否具有领域特异性的研究非常少，Maule 和 Franklin (2020)认为方差表征可能具有领域通用性。他们采用方差适应后效范

式，发现对方差大的色调刺激集合的适应会使被试低估之后呈现的朝向刺激的方差，这是视觉统计概要表征跨领域适应后效存在的首个证据。这一结果暗示了视觉方差表征是由一种领域通用机制编码的。

综上这些行为研究，不存在争议的是平均表征存在高—低水平领域间特异性的机制，而方差表征是否如此仍需探索，存在争议的是同一水平领域内不同特征间是否存在特异性机制。这些争议提示我们不仅需要行为研究进一步细致系统地考察，更需要对高和低不同水平以及同一水平不同特征间视觉统计概要表征背后的神经编码机制进行直接的考察，那将是具有重要开创性的工作。

3.2 统计概要表征不同统计指标之间的编码机制——平均表征和方差表征

对平均表征和方差表征适应后效的研究揭示出大脑内存在负责统计概要表征的特定神经编码(Corbett et al., 2012; Norman et al., 2015)。那么关于两者的加工机制是否相同，是探究视觉统计概要表征机制的重要问题。

一些研究发现平均表征和方差表征两者相互独立，无论是朝向还是颜色，对方差的适应后效不受平均值变化的影响(Maule & Franklin, 2020; Norman et al., 2015)，在双任务范式中，平均和方差表征这两个整体任务不会发生相互干扰，也说明两者可以独立估计(Khvostov & Utochkin, 2019)。另一方面，均值和方差辨别任务的成绩之间不存在显著的相关(Yang et al., 2018)。

但是，也有研究证据质疑均值和方差两者的独立性，在知觉启动研究中，只要变异性匹配(相同刺激维度的启动刺激和目标刺激的变异匹配)，被试可以很好地完成对目标的平均表征任务，表明变异性加工对平均表征存在影响(Michael et al., 2014)。反过来，平均表征也可以对方差表征存在影响(Tong et al., 2015)。类似地，在对朝向刺激的适应后效研究中，发现对方差的适应影响平均辨别，而对平均的适应影响之后的方差辨别(Jeong & Chong, 2020)。

4 视觉统计概要表征可能的神经机制

视觉统计概要表征是对繁复的视觉输入进行高效加工的过程，对其神经机制的探讨有利于我们理解视觉系统是怎样实现对均值等统计信息的计算过程。Haberman 和 Whitney (2012)以及 Whitney 等(2014)阐述了关于大脑进行统计概要表征的神经通路和表征过程的理论设想，认为视觉平均表征可能在视觉系统中的多个水平的层级上进行表征。如平均亮度、颜色和朝向，可能在早期的皮层甚至是皮层下区域进行表征，高水平的形状和面孔等可能在腹侧通路进行表征，而平均运动和位置则可能在背侧通路进行表征。平均表征过程则可能和神经信号汇集 (signal pooling) 有关。例如，当看到一组不同朝向的线段或者光栅时，朝向选择神经元(可能在 V1 中)被视觉刺激激活，来自部分或全部朝向选择神经元的活动被汇集起来，最终生成一个整体感知。

Corbett 等人(2012)关于圆点平均大小表征的适应后效研究揭示出大脑内存在负责统计概要表征的特定神经编码。适应大小不同的圆点集，然后判断测试圆的大小。在适应了一个大的平均尺寸的集合后，发现测试圆点被知觉得更小，反之亦然，即产生了知觉后效，作者认为平均大小似乎可作为一个特征维度在大脑中进行明确的神经表征。但是表征的脑区在哪里，从初级视觉皮层再到外侧枕叶编码客体大小的神经元群，皆有可能。行为研究揭示出对视觉统计信息的神经计算发生于双眼融合以及双眼抑制之后，即不早于初级视觉皮层。具体是，被试可以对分布于两眼的圆集合进行平均大小的判断，如果通过双眼竞争抑制某些项目，集合大小的平均表征会受到损害(Joo et al., 2009)。

但是直接考察统计概要表征的神经机制并不简单，一个主要的原因在于难以选择一个和集合的统计概要表征相匹配的对照条件，如果将单个刺激作为对照条件，集合和单个进行比较，刺激材料便不匹配，可能导致诱发的神经活动在一开始就存在差异。实际上，统计概要表征的研究道路也比较曲折，最初研究者通过行为实验推断统计概要表征背后可能的神经机制，后来有研究者考察了集合编码的神经机制，可以分为两类，一类是不包含统计概要信息的客体集合表征(类似于纹理，空间布局)，另一类是包含统计概要信息的客体集合表征，这类研究往往将多个刺激构成的集合与单个刺激进行比较，但是因为刺激数量不匹配，两者的差异并不能反映出纯粹的统计概要表征的神经机制。

4.1 集合表征中的空间分布和异常值

Cant 和 Xu 的课题组围绕海马旁回在集合表征中的功能进行了一系列的研究，发现当集合中的客体身份保持一致时(Cant & Xu, 2012)，甚至即使密度发生变化，海马旁回也会发生神经适应(Cant & Xu, 2015)，但是当集合中包含两类客体其相对密度发生变化时(Cant & Xu, 2015)以及集合中存在不同客体的异常值时(Cant & Xu, 2020)，适应程度都会受到影响。

Cant 和 Xu (2012)采用 fMRI 适应范式考察被试在加工物体集合时大脑的激活。发现当具有相同纹理统计的同一物体集合重复呈现时，即使是不同图片(局部特征不同)，包括海马旁回 (PPA) 在内的前腹内侧视皮层也会发生适应性变化。PPA 是加工场景和纹理的脑区，表明物体集合表征可能采用了和纹理统计表征相似的神经编码。

在另一项研究中，他们通过操控集合中个体的空间距离和相对比例探讨了绝对密度和相对密度对于集合表征的影响，发现当同一物体集合的密度发生改变时，PPA 也存在适应，适应量甚至和密度没有改变时相当。而当集合中包含两种物体，改变它们之间的比例，相应地两种物体各自的密度(相对密度)也发生变化，PPA 的神经适应减少。从而说明 PPA 对物体集合的加工更多依赖于物体比例等高水平视觉信息，而非间距或空间频率等低水平视觉信息(Cant & Xu, 2015)。

同样使用 fMRI 适应范式研究异常值在人脑中的表征。实验首先呈现一个包含某一客体的同质集合(如草莓)，然后呈现同一客体的同质集合(草莓)或大多数同一客体但带有少量不同客体即异常值的集合(草莓加少量西瓜)，被试需要判断两个集合是否相同，结果发现不同的异常值让与场景选择性相关的海马旁回 (PPA) 适应效应减弱。有趣的是，当首先呈现草莓集合，然后呈现西瓜集合或大多数西瓜加少量草莓，被试判断两个集合是否不同，发现匹配的异常值让 PPA 适应效应增强。这表明 PPA 等脑区可能在视觉感知过程中标记异常值，并在随后的行为决策中对异常值进行加权(Cant & Xu, 2020)。

Cant 和 Xu 的研究发现海马旁回对具有不同空间布局的同一集合会产生适应，但当集合的客体发生改变时，适应量会发生变化。具体为，海马对集合的表征不依赖空间布局这种统计属性，我们推测海马旁回对集合的表征可能依赖于概念，因此当相对密度或异常值对概念产生影响时，海马的适应也发生变化。

4.2 提取视觉统计概要表征时的集合编码

集合编码（Ensemble coding）指的是视觉系统对集合刺激的编码。我们主要围绕以下内容展开介绍，大脑中存在对集合属性进行整体加工的特定神经机制(Jia et al., 2022; Roberts et al., 2019)，而且集合编码和个体编码在

背侧通路和腹侧通路上出现了分离(Im et al., 2017)。光栅集合的神经机制和枕叶及顶叶有关，可能涉及了神经信号的汇集(signal pooling)(Tark et al., 2021)，也可能存在集合成员之间的相互作用(Jia et al., 2022)。

为了探究大脑是否存在处理集合整体感知的专门神经机制以及大脑如何计算集合的整体感知，研究者使用频率标记的脑电图法(SSVEP)来追踪大脑活动对周期性变化的圆环集合大小的整体表征，结果在顶枕电极上检测到了对集合整体大小的神经反应，接着他们使用时间响应函数(temporal response function, TRF)分离了成员个体大小以及个体间的相互作用(包含整体相互作用和局部相互作用，global interaction and local interaction)的神经反应，发现只有整体相互作用直接贡献了整体的大小感知。这些发现表明存在对集合大小进行整体加工的特定神经机制(Jia et al., 2022)。

为考察情绪面孔集合编码和情绪面孔个体编码的感知差异，研究者采用 fMRI 技术探索了两者激活脑区的分离。发现位于背侧通路的顶内沟和额上回参与情绪面孔的整体感知，而位于腹侧通路的梭状皮层参与情绪面孔的个体感知，同时这项研究中还发现了情绪面孔的整体编码具有大脑右半球偏侧化优势(Im et al., 2017)。2021 年他们使用 MEG 技术验证了之前的发现，即背侧通路参与了集合面孔的整体加工，腹侧通路对单个面孔进行识别与分辨，重要的是 MEG 揭示出背侧通路可以对集合面孔进行非常快速的整体编码。他们认为背侧通路可能依赖快速的大细胞通路的信息输入来形成对集合面孔情绪的整体表征(Im et al., 2021)。

研究者使用脑电 (EEG) 探索面孔身份统计概要表征的神经基础。实验中给被试呈现集合面孔或单个面孔，集合条件的 P1 波幅更小，N1 和 N2 的潜伏期更短。使用线性支持向量机(SVM)的多变量模式分析(MVPA)(将时间点和电极点纳入模型中)进行神经解码，发现神经信号不但可以区分不同的单张面孔，也可以区分平均身份不同的集合面孔，有趣的是，当两个集合的平均身份相同时，即使集合中的个体不同，神经信号也不能进行区分(边缘显著) (Roberts et al., 2019)。在这些研究者的另一篇文章中(实际上使用的面孔集合数据来自 Roberts 等(2019)的研究)，基于线性支持向量机和递归特征消除进行多变量特征选择，同样发现神经信号(在时间上和频率上)可以区分不同的面孔集合(Nemrodov et al., 2020)。

采用 fMRI 和逆向编码模型(inverted encoding model, IEM)，构建对光栅集合的平均朝向和成员个体朝向的选择性神经反应。研究者发现，虽然集合任务和个体任务的 BOLD 信号没有显著差异，但是 IEM 构建的神经反应却有所不同，在任务相关条件下，在枕叶和额顶叶都存在对平均朝向(集合任务中)以及个体朝向(个体任务中)的选择性神经反应，值得注意的是，在任务相关条件下对平均朝向的选择性反应在 V1 不显著，而在 V2 和 V3 显著，重要的是从 V1, V2 到 V3 存在显著的线性增长。该研究结果表明，视觉系统在多个层次上汇集的信号(pooled signal)形成了对整体感知的神经表征(Tark et al., 2021)。

研究者考察工作记忆表征是否像视觉系统的层级表征一样是结构化的，使用线条朝向的集合刺激，采用逆向编码模型(IEM)对 EEG 信号进行解码，结果支持结构化表征假设，在额中央区域观察到的静态编码不但可以表征简单的特征也可以表征抽象的集合平均值，并且与行为指标存在相关，在枕顶区观察到的动态编码和静态编码，受自上而下的任务需求调节(Oh et al., 2019)。

目前只有少量研究对方差的神经编码进行了初探。研究者发现对方差的适

应依赖于视网膜坐标(依赖于注视点位置)，而非空间视觉坐标。并且通过对一位只有左半球V1区未受损的脑损伤被试研究发现，其与正常被试相比，方差表征的正确率无明显差异，该被试双侧腹内侧枕颞皮质受损，右半球初级视觉皮层受损，双侧V2、V3、V4均有一部分受损，这表明可能方差表征发生在视觉系统的较早阶段，如初级视觉皮层(Norman et al., 2015)。总之，方差的神经编码涉及的脑区以及起源的时间早晚尚需大量研究进行探索。

5 研究展望

我们认为，大脑中同时存在两种统计概要表征机制，层级汇集的编码机制(signal pooling)和神经元集群编码机制(neural ensemble coding)，前者表征低水平的视觉刺激的统计属性，后者通过表征单个成员的神经元的同时激活来表征高水平的视觉刺激的统计属性。Whitney 等(2014)的理论模型认为平均表征过程可能和神经信号汇集有关，例如，当看到一组不同朝向的线段或者光栅时，朝向选择神经元被激活， 多个朝向选择神经元的活动汇集起来，最终生成整体集合感知。此理论得到实证研究证据的支持， 对平均朝向的选择性反应，从初级视觉皮层到纹外皮层存在显著的线性增长(Tark et al., 2021)。这一理论似乎很好地解释了低水平视觉刺激的统计概要表征，因为视觉皮层确实存在简单神经元以及接收并汇总简单神经元信息的下一级神经元即复杂神经元。然而对高水平视觉刺激如面孔刺激，当我们形成平均表征时，意味着大脑中存在高一级的神经元接收来自多个面孔的信息输入，这意味着大脑不但存在数量可观的表征单个面孔的神经元，也需要存在更加数目庞大的统计概要表征神经元。层级汇集的理论类似祖母细胞(Grandmother Cells) 理论同样将无法解决效率的问题。这时，统计表征依赖的可能是神经元集群编码，多个面孔的同时激活，形成了对集合整体的统计概要表征。因此对于统计概要表征，大脑中可能不但存在层级汇集的编码机制，同时也存在表征成员的神经元同时激活的集群编码机制。

尽管前文综述对统计概要表征的机制已经有了初步的理解，但仍然存在许多尚未解决或涉猎的问题需要未来研究的深入考察。

5.1 需要采用神经科学的技术手段，系统地考察统计概要表征的神经机制

首先，关于集合表征是依赖腹侧通路还是背侧通路，现有非常少量的研究并未取得一致的结论(Im et al., 2017; Im et al., 2021; Cant & Xu, 2020)，关于大细胞系统是否参与统计概要表征也存在不一致(Im et al., 2021; Lee & Chong, 2021)，未来研究需要关注腹侧通路和背侧通路以及大细胞系统和小细胞系统是否在统计概要表征过程中扮演不同的角色、是否具有不同的重要性，同时也需要关注自上而下的反馈信号以及同一水平内的相互作用信号是否在统计概要表征中扮演必不可少的重要角色，为了理清这些问题，也需要同时考虑刺激水平和任务的调节作用。

其次，需要关注统计概要表征的皮层起源以及加工的时间进程。由于初级视觉皮层的感受野非常小，往往不能覆盖集合中的多个刺激，似乎有理由相信统计概要表征发生于初级视觉皮层之后，然而此观点缺少神经科学研究的证据支持。来自行为实验的证据推论圆的大小的平均表征不早于初级视觉皮层(Joo et al., 2009)。对变异性的研究认为统计概要表征可能发生的时程更早，对一位纹外皮层和高级视觉区域存在损伤的患者的研究发现朝向的变异表征并未受损，说明朝向变异表征可能依赖初级视觉皮层(Norman et al., 2015)。对动物

的研究发现猫的外侧膝状体(lateral geniculate nucleus, LGN)可以对亮度的标准差进行特异性的反应，说明早期的视觉加工通路可以处理变异性信息(Bonin et al., 2006)。从动物实验到一个患者的脑损伤研究，再到行为研究，我们难以从中得出关于人类统计概要表征脑机制皮层起源的确定结论，未来研究在解决这一问题时也需要注意区分不同刺激水平可能造成的影响，同时也需要回答平均表征和变异性表征是否具有不同的神经机制。

最后，需要更直接地考察不同水平、不同指标(集中趋势和变异趋势)之间统计概要表征的特异性机制或通用性机制，到底是由特定脑区负责，还是共用一条加工通路，以及不同水平不同指标发生的时间进程是否具有一致性。对这些重要问题的揭示不能像前文所述的只停留在行为层面(Haberman, Brady & Alvarez, 2015)，未来需采用巧妙的实验设计，采用各种脑成像技术对这个问题进行直接的揭示。

5.2 统计概要表征是刻在基因里的还是可以受经验影响

统计概要表征是自动性比较高的心理过程，我们会很自然地预期基因的影响大于经验的影响。遗憾的是，少量关于文化差异的研究并不能解决这个问题。统计概要表征在自己种族面孔和他人种族面孔的比较上不存在显著差异，但是英国被试会对自己性别的面孔存在更强的平均倾向，而中国被试不存在这种倾向，可能因为视觉加工整体偏向不同(Peng et al., 2021)。一项关于 4-5 岁儿童对大小的统计概要表征研究中发现尽管这个时期的孩子视觉功能发展还尚未成熟(比如选择注意力、时空的注意分辨率以及视觉工作记忆容量等)，但他们已经能够对物体集合进行概要表征，识别平均大小而非集合中的某个个体，当然与成人相比，儿童表征的效率较低，这是因为随着年龄增长，视觉功能发展逐渐完善，统计概要表征的能力也有所提高(Sweeny et al., 2015)。此外，奖赏价值也许并不影响平均本身，而只影响统计概要表征的高水平的意识性表征(Dodgson & Raymond, 2020)。

5.3 集合编码和纹理表征的神经机制是否等同于统计概要表征的神经机制

很多考察神经机制的研究通过比较集合和个体的差异比较推论统计概要表征的神经机制，这种方法是存在问题的。集合和个体的神经活动差异或许包含了统计表征的脑机制，但也排除不了因为数量上的差别导致的神经活动的差异，即心理过程不单一纯粹。事实上，数量多少对神经活动存在影响，当集合数量增加时，N170的波幅增加(Puce et al., 2013)，甚至在初级视觉皮层集合的多个项目引起的C1波幅可以是其成员项目的线性加和(Chen et al., 2016)。而纹理表征和统计概要表征相比，似乎缺少了神经计算过程。未来的研究需要采用巧妙的方法分离出纯粹的统计概要表征的神经活动。

参考文献:

[1] 仝 可, 唐. 薇., 陈文锋, 傅小兰 . (2015). 统计概要表征的内容与机制. 心理科学进展, 23(10), 1723–1731. https://doi.org/10.3724/sp.J.1042.2015.01723

[2] Ariely, D. (2001). SEEING SETS:Representation by Statistical Properties. *PSYCHOLOGICAL SCIENCE*, *12*, 157-162. https://doi.org/10.1111/1467-9280.00327

[3] Attarha, M., & Moore, C. M. (2015). The perceptual processing capacity of summary statistics between and within feature dimensions. *Journal of Vision*, *15*(4), 9. https://doi.org/10.1167/15.4.9

[4] Banno, H., & Saiki, J. (2012). Calculation of the mean circle size does not circumvent the bottleneck of crowding. *Journal of Vision*, *12*(11). https://doi.org/10.1167/12.11.13

[5] Bonin, V., Mante, V., & Carandini, M. (2006). The statistical computation underlying

contrast gain control. *The Journal of Neuroscience*, *26*(23), 6346-6353. https://doi.org/10.1523/JNEUROSCI.0284-06.2006
[6] Cant, J. S., & Xu, Y. (2012). Object ensemble processing in human anterior-medial ventral visual cortex. *The Journal of Neuroscience*, *32*(22), 7685-7700. https://doi.org/10.1523/JNEUROSCI.3325-11.2012
[7] Cant, J. S., & Xu, Y. (2015). The Impact of Density and Ratio on Object-Ensemble Representation in Human Anterior-Medial Ventral Visual Cortex. *Cerebral Cortex*, *25*(11), 4226-4239. https://doi.org/10.1093/cercor/bhu145
[8] Cant, J. S., & Xu, Y. (2020). One bad apple spoils the whole bushel: The neural basis of outlier processing. *Neuroimage*, *211*, 116629. https://doi.org/10.1016/j.neuroimage.2020.116629
[9] Chen, J., Yu, Q., Zhu, Z., Peng, Y., & Fang, F. (2016). Spatial summation revealed in the earliest visual evoked component C1 and the effect of attention on its linearity. *J Neurophysiol*, *115*(1), 500-509. https://doi.org/10.1152/jn.00044.2015
[10] Corbett, J. E., Wurnitsch, N., Schwartz, A., & Whitney, D. (2012). An aftereffect of adaptation to mean size. *Visual Cognition*, *20*(2). https://doi.org/10.1080/13506285.2012.657261
[11] Davis, E. E., Matthews, C. M., & Mondloch, C. J. (2021). Ensemble coding of facial identity is not refined by experience: Evidence from other-race and inverted faces. *British Journal of Psychology*, *112*(1), 265-281. https://doi.org/10.1111/bjop.12457
[12] de Fockert, J., & Wolfenstein, C. (2009). Rapid extraction of mean identity from sets of faces. *Quarterly journal of experimental psychology*, *62*(9), 1716-1722. https://doi.org/10.1080/17470210902811249
[13] Dodgson, D. B., & Raymond, J. E. (2020). Value associations bias ensemble perception. *Attention, Perception, & Psychophysics*, *82*(1), 109-117. https://doi.org/10.3758/s13414-019-01744-1
[14] Epstein, M. L., & Emmanouil, T. A. (2021). Ensemble Statistics Can Be Available before Individual Item Properties: Electroencephalography Evidence Using the Oddball Paradigm. *Journal of Cognitive Neuroscience*, *33*(6), 1056-1068. https://doi.org/10.1162/jocn_a_01704
[15] Fischer, J., & Whitney, D. (2011). Object-level visual information gets through the bottleneck of crowding. *J Neurophysiol*, *106*(3), 1389-1398. https://doi.org/10.1152/jn.00904.2010
[16] Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2009). Positional averaging explains crowding with letter-like stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(31), 13130 – 13135. https://doi.org/https://doi.org/10.1073/pnas.0901352106
[17] Haberman, J., & Whitney, D. (2007a). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*, *17*(17), R751-753. https://doi.org/10.1016/j.cub.2007.06.039
[18] Haberman, J., & Whitney, D. (2007b). *Supplemental Data: Rapid extraction of mean emotion and gender from sets of faces*. 5.
[19] Haberman, J., Harp, T., & Whitney, D. (2009). Averaging facial expression over time. *Journal of Vision*, *9*(11), 1 – 1. https://doi.org/10.1167/9.11.1
[20] Haberman, J., & Whitney, D. (2009). Seeing the mean: ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 718-734. https://doi.org/10.1037/a0013899
[21] Haberman, J., & Whitney, D. (2012). Ensemble perception: Summarizing the scene and broadening the limits of visual processing. *From perception to consciousness: Searching with Anne Treisman*, 339-349.
[22] Haberman, J., Brady, T. F., & Alvarez, G. A. (2015). Individual differences in ensemble perception reveal multiple, independent levels of ensemble representation. *Journal of Experimental Psychology: General*, *144*(2), 432-446. https://doi.org/10.1037/xge0000053
[23] Haberman, J., Lee, P., & Whitney, D. (2015). Mixed emotions: Sensitivity to facial variance in a crowd of faces. *Journal of Vision*, *15*(4), 16. https://doi.org/10.1167/15.4.16
[24] Haberman, J. M., & Ulrich, L. (2019). Precise Ensemble Face Representation Given Incomplete Visual Input. *i-Perception*, *10*(1), 2041669518819014. https://doi.org/10.1177/2041669518819014
[25] Haberman, J., & Suresh, S. (2021). Ensemble size judgments account for size constancy. *Attention, Perception, & Psychophysics*, *83*(3), 925 – 933. https://doi.org/10.3758/s13414-020-02144-6

[26] Haxby, J. V., Hoffman, E., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223-233. https://doi.org/10.1016/S1364-6613(00)01482-0
[27] Im, H. Y., & Chong, S. C. (2009). Computation of mean size is based on perceived size. *Attention, Perception, & Psychophysics*, *71*(2), 375-384. https://doi.org/10.3758/APP.71.2.375
[28] Im, H. Y., Albohn, D. N., Steiner, T. G., Cushing, C. A., Adams, R. B., Jr., & Kveraga, K. (2017). Differential hemispheric and visual stream contributions to ensemble coding of crowd emotion. *Nature Human Behaviour*, *1*, 828-842. https://doi.org/10.1038/s41562-017-0225-z
[29] Im, H. Y., Cushing, C. A., Ward, N., & Kveraga, K. (2021). Differential neurodynamics and connectivity in the dorsal and ventral visual pathways during perception of emotional crowds and individuals: a MEG study. *Cognitive, Affective, & Behavioral Neuroscience*, *21*(4), 776-792. https://doi.org/10.3758/s13415-021-00880-2
[30] Jeong, J., & Chong, S. C. (2020). Adaptation to mean and variance: Interrelationships between mean and variance representations in orientation perception. *Vision Research*, *167*, 46-53. https://doi.org/10.1016/j.visres.2020.01.002
[31] Jia, J., Wang, T., Chen, S., Ding, N., & Fang, F. (2022). Ensemble size perception: Its neural signature and the role of global interaction over individual items. *Neuropsychologia*, *173*, 108290. https://doi.org/10.1016/j.neuropsychologia.2022.108290
[32] Ji, L., Rossi, V., & Pourtois, G. (2018). Mean emotion from multiple facial expressions can be extracted with limited attention: Evidence from visual ERPs. *Neuropsychologia*, *111*, 92-102. https://doi.org/10.1016/j.neuropsychologia.2018.01.022
[33] Joo, S. J., Shin, K., Chong, S. C., & Blake, R. (2009). On the nature of the stimulus information necessary for estimating mean size of visual arrays. *Journal of Vision*, *9*(9), 7 1-12. https://doi.org/10.1167/9.9.7
[34] Kacin, M., Gauthier, I., & Cha, O. (2021). Ensemble coding of average length and average orientation are correlated. *Vision Research*, *187*, 94-101. https://doi.org/10.1016/j.visres.2021.04.010
[35] Khvostov, V. A., & Utochkin, I. S. (2019). Independent and parallel visual processing of ensemble statistics: Evidence from dual tasks. *Journal of Vision*, *19*(9), 3. https://doi.org/10.1167/19.9.3
[36] Lee, J., & Chong, S. C. (2021). Quality of average representation can be enhanced by refined individual items. *Attention, Perception, & Psychophysics*, *83*(3), 970-981. https://doi.org/10.3758/s13414-020-02139-3
[37] Li, H., Ji, L., Tong, K., Ren, N., Chen, W., Liu, C. H., & Fu, X. (2016). Processing of Individual Items during Ensemble Coding of Facial Expressions. *Frontiers in Psychology*, *7*, 1332. https://doi.org/10.3389/fpsyg.2016.01332
[38] Maule, J., & Franklin, A. (2020). Adaptation to variance generalizes across visual domains. *Journal of Experimental Psychology: General*, *149*(4), 662-675. https://doi.org/10.1037/xge0000678
[39] Michael, E., de Gardelle, V., & Summerfield, C. (2014). Priming by the variability of visual information. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(21), 7873-7878. https://doi.org/10.1073/pnas.1308674111
[40] Nemrodov, D., Ling, S., Nudnou, I., Roberts, T., Cant, J. S., Lee, A. C. H., & Nestor, A. (2020). A multivariate investigation of visual word, face, and ensemble processing: Perspectives from EEG-based decoding and feature selection. *Psychophysiology*, *57*(3), e13511. https://doi.org/10.1111/psyp.13511
[41] Nguyen, T. T. N., Vuong, Q. C., Mather, G., & Thornton, I. M. (2021). Ensemble coding of crowd speed using biological motion. *Attention, Perception, & Psychophysics*, *83*(3), 1014-1035. https://doi.org/10.3758/s13414-020-02163-3
[42] Norman, L. J., Heywood, C. A., & Kentridge, R. W. (2015). Direct encoding of orientation variance in the visual system. *Journal of Vision*, *15*(4), 3. https://doi.org/10.1167/15.4.3
[43] Oh, B. I., Kim, Y. J., & Kang, M. S. (2019). Ensemble representations reveal distinct neural coding of visual working memory. *Nature Communications*, *10*(1), 5665. https://doi.org/10.1038/s41467-019-13592-6
[44] Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, *4*(7), 739-744. https://doi.org/https://doi.org/10.1038/89532

[45] Peng, S., Zhang, L., Xu, R., Liu, C. H., Chen, W., & Hu, P. (2019). Self-Construal Priming Modulates Ensemble Perception of Multiple-Face Identities. *Frontiers in Psychology*, *10*, 1096. https://doi.org/10.3389/fpsyg.2019.01096
[46] Peng, S., Liu, C. H., & Hu, P. (2021). Effects of subjective similarity and culture on ensemble perception of faces. *Attention, Perception, & Psychophysics*, *83*(3), 1070-1079. https://doi.org/10.3758/s13414-020-02133-9
[47] Phillips, L. T., Slepian, M. L., & Hughes, B. L. (2018). Perceiving groups: The people perception of diversity and hierarchy. *Journal of Personality and Social Psychology*, *114*(5), 766-785. https://doi.org/10.1037/pspi0000120
[48] Puce, A., McNeely, M. E., Berrebi, M. E., Thompson, J. C., Hardee, J., & Brefczynski-Lewis, J. (2013). Multiple faces elicit augmented neural activity. *Frontiers in Human Neuroscience*, *7*, 282. https://doi.org/10.3389/fnhum.2013.00282
[49] Rajendran, S., Maule, J., Franklin, A., & Webster, M. A. (2021). Ensemble coding of color and luminance contrast. *Attention, Perception, & Psychophysics*, *83*(3), 911-924. https://doi.org/10.3758/s13414-020-02136-6
[50] Roberts, T., Cant, J. S., & Nestor, A. (2019). Elucidating the Neural Representation and the Processing Dynamics of Face Ensembles. *The Journal of Neuroscience*, *39*(39), 7737-7747. https://doi.org/10.1523/JNEUROSCI.0471-19.2019
[51] Sama, M. A., Nestor, A., & Cant, J. S. (2019). Independence of viewpoint and identity in face ensemble processing. *Journal of Vision*, *19*(5), 2. https://doi.org/10.1167/19.5.2
[52] Sun, J., & Chong, S. C. (2020). Power of averaging: Noise reduction by ensemble coding of multiple faces. *Journal of Experimental Psychology: General*, *149*(3), 550-563. https://doi.org/10.1037/xge0000667
[53] Sun, P., Chu, V., & Sperling, G. (2021). Multiple concurrent centroid judgments imply multiple within-group salience maps. *Attention, Perception, & Psychophysics*, *83*(3), 934 - 955. https://doi.org/10.3758/s13414-020-02197-7
[54] Sweeny, T. D., Haroz, S., & Whitney, D. (2013). Perceiving group behavior: Sensitive ensemble coding mechanisms for biological motion of human crowds. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 329 - 337. https://doi.org/10.1037/a0028712
[55] Sweeny, T. D., Wurnitsch, N., Gopnik, A., & Whitney, D. (2015). Ensemble perception of size in 4-5-year-old children. *Developmental science*, *18*(4), 556-568. https://doi.org/10.1111/desc.12239
[56] Tark, K. J., Kang, M. S., Chong, S. C., & Shim, W. M. (2021). Neural representations of ensemble coding in the occipital and parietal cortices. *Neuroimage*, *245*, 118680. https://doi.org/10.1016/j.neuroimage.2021.118680
[57] Tiurina, N. A., & Utochkin, I. S. (2019). Ensemble perception in depth: Correct size-distance rescaling of multiple objects before averaging. *Journal of Experimental Psychology: General*, *148*(4), 728-738. https://doi.org/10.1037/xge0000485
[58] Tokita, M., Ueda, S., & Ishiguchi, A. (2016). Evidence for a Global Sampling Process in Extraction of Summary Statistics of Item Sizes in a Set. *Frontiers in Psychology*, *7*, 711. https://doi.org/10.3389/fpsyg.2016.00711
[59] Tong, K., Ji, L., Chen, W., & Fu, X. (2015). Unstable mean context causes sensitivity loss and biased estimation of variability. *Journal of Vision*, *15*(4), 15. https://doi.org/10.1167/15.4.15
[60] Ward, E. J., Bear, A., & Scholl, B. J. (2016). Can you perceive ensembles without perceiving individuals?: The role of statistical perception in determining whether awareness overflows access. *Cognition*, *152*, 78-86. https://doi.org/10.1016/j.cognition.2016.01.010
[61] Wardle, S. G., Bex, P. J., Cass, J., & Alais, D. (2012). Stereoacuity in the periphery is limited by internal noise. *Journal of Vision*, *12*(6), 12 - 12. https://doi.org/10.1167/12.6.12
[62] Watamaniuk, S. N. J., & Duchon, A. (1992). The human visual system averages speed information. *Vision Research*, *32*(5), 931 - 941. https://doi.org/10.1016/0042-6989(92)90036-I
[63] Whitney, D., Haberman, J., & Sweeny, T. D. (2014). 49 From textures to crowds: multiple levels of summary statistical perception. In J. S. Werner, L. M. Chalupa & M. E. Burns (Eds.), *The new visual neurosciences* (pp. 695 - 710). Cambridge, MA: MIT Press.
[64] Whitney, D., & Yamanashi Leib, A. (2018). Ensemble Perception. *Annu Rev Psychol*, *69*, 105-129. https://doi.org/10.1146/annurev-psych-010416-044232

[65] Yang, Y., Tokita, M., & Ishiguchi, A. (2018). Is There a Common Summary Statistical Process for Representing the Mean and Variance? A Study Using Illustrations of Familiar Items. *i-Perception*, *9*(1), 2041669517747297. https://doi.org/10.1177/2041669517747297
[66] Ying, H., & Xu, H. (2017). Adaptation reveals that facial expression averaging occurs during rapid serial presentation. *Journal of Vision*, *17*(1), 15. https://doi.org/10.1167/17.1.15
[67] Yoruk, H., & Boduroglu, A. (2020). Feature-specificity in visual statistical summary processing. *Attention, Perception, & Psychophysics,* *82*(2), 852-864. https://doi.org/10.3758/s13414-019-01942-x

通讯作者:张秀玲 E-mail: zhangxl556@nenu.edu.cn
蒋毅 E-mail: yijiang@psych.ac.cn
张帆 E-mail: yffs9762@163.com

作者贡献声明:

张秀玲：负责论文撰写修改、论文审核及终版修订、文献整理、投稿支持及协商
蒋毅： 负责论文监督及审核
张帆：负责论文初稿撰写、论文内容整理补充、文献收集整理
葛明晓：负责补充文献资源及论文内容
李思嘉：负责补充文献资源及论文内容、参考文献整理